DOCUMENT RESUME

ED 452 826 IR 020 663

AUTHOR Shakeshaft, Charol
TITLE Measurement Issues with Instructional and Home Learning Technologies.
PUB DATE 1999-00-00
NOTE 9p.; In: The Secretary's Conference on Educational Technology, 1999: Evaluating the Effectiveness of Technology. [Proceedings] (Washington, D.C., July 12-13, 1999); see IR 020 654.
AVAILABLE FROM For full text: http://www.ed.gov/technology/1999/whitepapers.
PUB TYPE Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; Computer Assisted Instruction; *Computer Uses in Education; *Educational Research; *Educational Technology; *Evaluation Criteria; Measurement; Measurement Techniques; Program Evaluation; Technology Integration
IDENTIFIERS Technology Implementation

ABSTRACT

Evaluating the effects of technology use provokes the same evaluation challenges as does any other program intervention. The issues addressed in this paper are based on experience in evaluating the achievement effects of specific technology implementations. Five previous studies of technology use and student achievement are identified in a table, each with purpose, sample/setting, method and data collection, and findings. The remainder of the paper deals with the following topics: determining the technology "input" to be measured; how to measure the amount of student use or exposure; and how to know if technology works, by measuring dependent variables. An appendix lists 10 practical Frequently Asked Questions (FAQs) about measuring information technology effects. Selected sources on measurement of instructional technology are provided. (AEF)

Reproductions supplied by EDRS are the best that can be made from the original document.

ED 452 826

Measurement Issues with Instructional and Home Learning Technologies

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

By: Charol Shakeshaft

IR020663

BEST COPY AVAILABLE

Measurement Issues with Instructional and Home Learning Technologies

I. INTRODUCTION
II. WHAT ARE WE MEASURING?
III. MEASURING USE OR EXPOSURE
IV. HOW DO WE KNOW IF TECHNOLOGY WORKS? MEASURING THE DEPENDENT VARIABLE(S)
V. CONCLUSION

APPENDIX

Ten Practical FAQ's (Frequently Asked Questions) about measuring IT effects

Selected Sources on Measurement of Instructional Technology

I. INTRODUCTION

Evaluating the effects of technology use provokes the same evaluation challenges as does any other program intervention. The issues that I address in this paper are based upon my experience in evaluating the achievement effects of specific technology implementations. The five studies that have offered me the largest learning laboratory are listed in Table 1. Each required a careful description of the technology to be studied, a measure of how much students used the technology, and a measure of achievement gains.

As Mann has pointed out in "Documenting the Effects of Instructional Technology: A Fly-Over of Policy Questions", a variety of stakeholders are beginning to ask questions about technology use in schools. Many of these questions go no further than "Does technology work?" Or, "Does technology use improve student achievement?"; "Is technology in schools worth the money it costs?"; "Are there benefits to students beyond achievement?"

Table 1:
Studies of Technology Use and Student Achievement

Study	Purpose	Sample/ Setting	Method and Data Collection	Findings

The Cyberspace Regionalization Project: "Virtual Desegregation"	Can audio-visual telecommunications be used to bridge gaps of geography, race and social class?	650 9th grade students in two high schools: one upper income with Caucasian students; one lower income with students of African descent.	Four year study. Interviews, surveys, annual pre-post administration of a Racial Attitude Assessment Instrument, administrative data transfer. Four year data collection.	Study in progress. However, baseline data collected in the Fall of 1998 reveal gaps in interracial contact and significant variation in racial attitude scales.
Lightspan Achieve Now and the Home-School Connection: Adams 50, Westminster, CO.	Does implementation of a game-like, CD-ROM-based, K-6 curriculum launched at school and used at home with families improve student achievement in math and language arts?	6 elementary schools; 2,000+ students and 55 teachers in grades 2-5.	Three year study of 3 elementary schools using Lightspan compared with 3 not using Lightspan. Annual pre-post Terra Nova data, district reading test scores, Colorado test scores. Observations in classrooms. Interviews with parents, teachers, and students four times each year. Learning Combination Inventory. On-line data collection.	After one-year of implementation, the students in 3 treatment schools surpassed students in the control schools and significantly outperformed them on the CTB-Terra Nova (Reading and Math).

Read 180 Scholastic, Inc. (National, urban settings)	Can a CD Rom interactive basic skills curriculum remediate prior deficiencies for early adolescents who are 4 or more grades behind in achievement?	Random assignment of 1,400 6th and 7th grade students to Read 180 and control classrooms in 7 big city school districts (Chicago, Dallas, Miami- Dade, Houston, Atlanta, San Francisco, and Boston)	Two year pre-post measures in Stanford 9 Language Arts subtests in Read 180 and control classrooms. Self efficacy, discipline, achievement in other subject areas, and attitude toward school are examined.	Data collection begins September 1999.
Technology Impact Study in the Schools of the Mohawk Region, New York State	What is the impact on student achievement associated with a $14.1 million investment in educational technology?	55 school districts, 4,041 teachers, 1,722 students, 159 principals, 41 superintendents	Teacher survey, principal survey, administrative data transfer of New York State PEP and Regents test scores	For the schools that had the most technology and training for teachers, the average increase in the percentage of students who to ok and passed the Math Regents Exam was 7.5; the average increase for the English Regents Exam was 8.8.
West Virginia's Basic Skills/Computer Education (BS/CE) Program	What effect does a $70 million statewide comprehensive instructional technology program have on student achievement?	18 elementary schools, 950 fifth grade students, teachers and principals in all the schools	Teachers survey Principal survey Student survey Observations Interviews with principals, teachers, and students Stanford 9 data for two years	A BS/CE technology regression model accounts for 11% of the total variance and 33% of the within school variance in the one-year basic skills achievement gain scores.

II. WHAT ARE WE MEASURING?

Many of the program administrators responsible for IT have not thought through the questions they want answered by documentation research, nor can they be expected to since operational responsibilities often preempt evaluation. Part of the job of the evaluator is crafting work that serves the needs of the stakeholders: Is this an evaluation for re-funding? For use in curriculum refinement? For analysis of classroom instruction? For public relations? For all of these?

Because stakeholder needs are not always clear, the first measurement challenge is to determine the technology "input" to be examined. Technology is lots of things: computers, CD-ROM and videodisc players, networked applications. If we focus on computers, it generally is not the use of the computer per se that is of interest, but rather a specific use, especially particular software.

For most readers of this paper, the "what is the technology" question will seem elementary. However, my experience has been that many stakeholders -- particularly school administrators, school board members, and legislators -- expect that if hardware is purchased, then improved achievement should follow. A common situation we have faced is being asked to determine achievement gains in schools where computers and word processing software are purchased. The notion that doing anything on a computer should lead to (any) achievement gains is widespread. *(We were once asked to measure the math achievement impact of having provided Corel's WordPerfect word-processing software to all the elementary teachers of a district!)* Therefore, identifying what technology use is being analyzed is a first step, and a step I would not bother to relate had I not learned the hard way that identifying the technology to be measured requires a considerable amount of interaction with stakeholders.

Is the technology question really a focus on the teaching efficacy of a particular software that students are using? If so, is there a relationship between the software design characteristics and student achievement? Do any of the following make a difference: instructional control, feedback, objectives and advance organizers, cognitive strategies, conceptual change strategies, scaffolding of learning support, still and animated graphics, dynamic visualization, video, navigational technique, text and story content, game context and visual metaphor fantasy context, Window presentation styles?

Or, is the question about multiple sites for technology use? The home? The school? Both? And if so, how much of what interaction in which site is related to achievement?

Do different technologies result in different kinds of achievement? For instance, do telecommunication distance learning technologies such as access to online resources, document exchange and discussion, or professional development on-line improve student achievement? If they do, is this be a direct relationship? How would we isolate these uses while examining student achievement?

It is easy to see how an initially simple question like, "What is the relationship between technology use and student achievement?" blossoms into refinements and further definitions. Carefully defining the technology to be studied then takes us to the next step.

III. MEASURING USE OR EXPOSURE

Just because technology is present does not mean that the students are using it. How do we measure the intensivity of student use?

We faced this question in every study we have done. We have used observations, file server records,

student reports, parent reports (thousands of telephone interviews, each logged and coded), teacher reports, and on-site observations. Because it isn't feasible to shadow every student every day, observational data, although probably both reliable and valid, is not often feasible. Metering and file server records, although able to record time on the computer or software, are not available in most schools. The next level of data is self report data from students, which can be verified by teachers and parents. If we are examining the relationships between the use of some technology and student achievement, we do sampled surveys of use. We ask students, teachers, and parents about the previous day or week's activity. We use e-mail, web-site, telephone, face-to-face, and paper and pencil surveys to document student use.

Not surprisingly, filling out surveys is not a priority for many educators, whether they are sent by e-mail, snail mail, or over telephone lines, but we have always had excellent cooperation that easily exceeds the minimum standards for sample size and response. Student reports of their own behavior tend to be more accurate than parent or teacher responses, although children younger than fifth grade often have difficulty estimating time. Teachers are usually able to tell us how much in-class time that students spend on the computer, although it often depends on which day, which class, and which student. Teacher reports are aggregate reports, while student reports are specific to the individual student.

Because student use (at least in schools) is related to teacher use of and comfort with technology, we include in the description of the technology the amount of teacher professional development and integration into the curriculum. We ask teachers and administrators about use. We examine teacher professional development participation, both in school and out of school, formal and informal. Self reports of technology literacy, faculty meeting agendas, lesson plans, and observations all help to describe what the teacher knows about technology, how comfortable the teacher is with technology, and how and how often the teacher is able to integrate technology into the curriculum.

IV. HOW DO WE KNOW IF TECHNOLOGY WORKS? MEASURING THE DEPENDENT VARIABLE(S)

While this paper is about measurement issues and student achievement, there are worthy reasons to use technology beyond bottom-line achievement. We have examined technology use and self efficacy, attitude about school, attendance, and discipline.

However, to understand the relationship between technology use and student achievement, we are most comfortable with examining gains in individual student achievement that would be reasonably expected because of the technology. Thus, we don't expect that time using music composition software would accelerate student learning in biology. The measures used must relate to the expectations of the technology.

We use the same data that schools use to determine achievement, even when we might not think it is the best form of measurement. We use these data because that is how the districts and their superordinate jurisdictions measure achievement. While we can argue that most achievement tests do not accurately or fully explain what students learn, the reality is that achievement data is often the best we have.

Thus, we often rely upon gain scores from September to May on norm referenced tests such as the Stanford 9, the Iowa Tests of Basic Skills, or CTB-Terra Nova. Since most districts don't test twice a year, this usually requires some negotiation. However, the result is that we have individual student gain

scores to relate to the individual student use measures.

Additionally, we use grade, teacher developed tests, state achievement tests, district achievement tests, and authentic displays of student work. The more types of data, the better the understanding.

V. CONCLUSION

If you look across the measurement literature (and Jay Sivin-Kachalan and Ellen Bialo have, see sources below), you will find different methods to study different combinations of different interventions. It is hard to make those disparate studies add up in a way that compels belief. In part, that is the nature of decentralized science in a democracy. Still, we would like to see a short list of preferred evaluation methods or models, each for example, with two alternative methods for different intervention niches like early childhood literacy or gender studies of literacy applications delivered on the Internet. We would like to see those models developed and recommended (or even encouraged) by funding agencies. That way, at least some of what we do would add up in a more direct fashion than has so far been the case.

Measuring technology outcomes is undeniably messy and imperfect. It is also important for the practice-improving signals that can be developed even from this sometimes frustrating enterprise. It may also be helpful to recognize that just as instructional technology continues to evolve and to improve, so does our ability to document inputs and measure effects.

About the Author: Charol Shakeshaft, Ph.D., is professor in the Department of Administration and Policy Studies, School of Education, Hofstra University, Hempstead NY 11590. An internationally recognized expert in gender studies and women's leadership in school administration, Professor Shakeshaft's new book is In Loco Parentis: Sexual Abuse in the Schools (San Francisco, Josey-Bass, in press). Dr. Shakeshaft is a Managing Director of Interactive, Inc., 326 New York Avenue, Huntington, NY 11743-3360: p 516 5470464: f 516 547 0465.

APPENDIX

Ten Practical FAQ's (Frequently Asked Questions) about measuring IT effects

1. **Q**: It is too early to expect results. **A**: It is always too early but if there is a partial implementation (which is almost always the case anyway) then we need sensitive measures and an expectation of probably faint signals of effect.
2. **Q**: Instructional Technology wasn't the only thing we did. We changed textbooks, moved to a house plan, etc. **A**: Good, there are no single answers, not even technology. If the documentation plan calls for measuring the different dimensions of all the things that were going on, then regression analysis will allow testing for differences in the strength of relationships between different input clusters and outcome measures.
3. **Q**: We changed tests two years ago. Can we still look for effects? **A**: Everybody changes tests and that is more of an inconvenience to the analyst than a barrier to inquiry. The whole point of nationally normed tests is to facilitate comparison.
4. **Q**: We keep changing and replacing both hardware and software. How can we know which version of what makes a difference? **A**: That's an excellent question. We all need to do a better job of

keeping track of what hardware/software experiences which kids had.

5. **Q**: Doesn't it take thousands of cases to do good research? Our district(school) isn't that big? **A**: With well constructed samples, it is possible to generalize to the population from surprisingly small numbers of respondents. Selecting those sampling dimensions (and getting access to schools, teachers and children) is one of the places where the client organizations can be helpful.
6. **Q**: How can you say for sure that IT "caused test score gains"? **A**: Strictly speaking, none of us can make that claim on the research designs that are practically feasible. But social science research is seldom if ever causal. One way or the other, decision makers have to commit their organizations. We try to help with the best data from the most powerful designs we can get.
7. **Q**: If somebody outside the school district pays for the study, then it isn't objective. **A**: We do lots of studies paid for by third parties. The question is not, who paid for it, but how was it done. We always report our methods (sample, data collection instruments and techniques, analysis procedures) and we make that publicly available. If everyone follows the rules of science and if the study followed those rules, then the objectivity is there regardless of the auspices.
8. **Q**: It takes millions of dollars to do good research. **A**: Research that ends up with compelling results is sometimes costly. But we find that districts and schools will help with data collection, they do part of the work of mailing, they critique procedures and generally share costs to make things feasible at modest prices.
9. **Q**: The most important question is, does IT change the act of teaching? How can you find that out? **A**: We believe in multiple methods. That's why most of our work is quantitative/qualitative (or vice versa) in successive waves. Lots of people think that IT can help teachers use more constructivist methods and we have been developing and refining item banks to measure just that---the shift from instructivist to constructivist.
10. **Q**: Evaluations are always ignored. **A**: Some are. It depends on how directly (and simply) the reports and the underlying data speak to the policy issues. And also on the patience of the policy makers and of the measurement people.

Selected Sources on Measurement of Instructional Technology

- International Society for Technology in Education (ISTE) (1998). *National Educational Technology Standards for Students.* Eugene, Or. (funded by the National Aeronautics and Space Administration (NASA) in consultation with the U.S. Dept. of Education; the Milken Exchange on Education Technology; and Apple Computer, Inc.) (www.iste.org).
- The CEO Forum on Education and Technology (1997). *School Technology and Readiness Report: From Pillars to Progress.* Washington, D.C. (www.ceoforum.org).
- Milken Exchange on Education Technology. (1998). *Seven Dimensions for Gauging Progress.* Santa Monica, CA. (www.mff.org).
- Sivin-Kachala, J. & Bialo, E.R. (1999) (For the Software & Information Industry Association). *1999 Research Report on the Effectiveness of Technology in Schools.* Washington, D.C. (www.siia.net).

-####-